

要旨

クラウドコンピューティングにおける データセンタ間連携による可用性向上法の研究

阿部 逸人

近年，クラウドコンピューティングが注目されている．クラウドコンピューティングとは，ネットワーク上にあるサーバが提供するサービスをサーバ群を意識することなく利用可能なコンピューティングモデルのことを指す．クラウドコンピューティングを使って提供されるサービスをクラウドサービスと呼び，その市場は拡大傾向にある．

このクラウドサービスを提供するために，多数のサーバを内包するデータセンタが活用されている．今後クラウドサービスが増加し，1つのデータセンタ内にあるサービスが増加した場合，データセンタ内のリソースが減少する．その結果，1つのサービスに対して大量のリクエストが来た場合，そのデータセンタ内だけでは処理しきれない可能性がある．

本論文では，連携用サーバを使ったデータセンタ間での連携を行なう方式を提案した．提案方式は，1つのデータセンタ内で対処しきれないほどのリクエストが送られて来ることを連携用サーバが予測する．そして条件に合う他のデータセンタを探索し，サービスコピーを行う．その後コピー先のとコピー元のデータセンタ間でリクエストを分散させて処理を行う．これにより1つのデータセンタで処理できる限界以上のリクエストに対しても処理することができる．また1つのデータセンタへの負荷を一定以下に抑えることができる．

有効性を示すために，Visual SLAM によるシミュレーション実験を行い，提案方式の性能評価を行った．提案方式を適用した結果，1つのデータセンタにおける同時処理可能な負荷の120%分の処理を行うことが可能となることを確認し，なおかつ1つのデータセンタでの処理量は同時処理可能な負荷の66%に抑えることができた．また，処理負荷増加の予測と

サービスコピー時間を一定にした時に，1つのデータセンターにおける同時処理可能な負荷が5秒に1%ずつ増えた場合でも対応できることを確認した．

キーワード クラウドコンピューティング，データセンター間連携，可用性

Abstract

A study of availability improvement method by the cooperation among data centers in cloud computing

HAYATO Abe

Recently, cloud computing is attracting attention. Cloud computing is the one that the resource and service in which users existed on the network were made available through the network. The service provided by using cloud computing is called cloud service. Data center containing a lot of servers is utilized to offer this cloud service. Service in one data center will increase when cloud service increases in the future. And resources in the data center decrease. As a result, there is a possibility not to be able to process in the data center alone when a large amount of request comes to one service.

Then, I proposed the method to do the cooperation between the data centers that used the server for cooperation. The server for cooperation predicts that the request that cannot deal in one data center has been sent. And the server search for data center in accord with a condition and copy the service. The server do distributed processing of a request afterwards between the data center of copying and the origin of copy. The server can handle it for the requests more than the limit that I can handle by this in one data center. In addition, it can hold down load to one data center to the constant following. Then I performed the performance evaluation of the suggestion method with Visual SLAM. The server was confirmed to be able to process 120% of the load of a possible simultaneous processing in one data center worth. And the server was able to hold down the throughput in one data center to 66% of the load that the server could handle at the same time. In addition, when a prediction and service copy time for processing

load increase are constant,even if the load that I can handle at the same time increases by 1% for five seconds, I can cope.

key words cloud computing, the cooperation among data centers, availability